

CLAIMS

What is claimed is:

- 1 1. A method comprising:
 - 2 receiving a first utterance from an intended talker at an integrated
 - 3 speech and speaker recognition system;
 - 4 generating a voice characteristic model for the intended talker;
 - 5 receiving a second utterance from the intended talker in a noisy
 - 6 environment at the speaker recognition system;
 - 7 processing a portion of speech associated with the second utterance,
 - 8 wherein processing comprises,
 - 9 computing a speaker verification score with the voice characteristic
 - 10 model associated with the portion of speech,
 - 11 computing a speech recognition score associated with the portion
 - 12 of speech, and
 - 13 generating a combined score by combining the speaker verification
 - 14 score and the speech recognition score; and
 - 15 selecting a best hypothesis associated with the second utterance and
 - 16 based upon the combined score.
- 1 2. The method of claim 1, wherein the portion of speech includes a word, a
- 2 sentence, a syllable, or a frame.

3. The method of claim 1, wherein processing further comprises altering a search path in a Viterbi search used by a speech recognizer.

1 4. The method of claim 1, wherein identifying a intended talker comprises using
2 hotword speech recognition to identify the intended talker.

1 5. The method of claim 1, wherein the noisy environment includes additional
2 speakers, music, stationary and non-stationary noise.

1 6. The method of claim 1, wherein the voice characteristic model includes a
2 voice print, personal profile and linguistic characteristics.

1 7. A system comprising:

2 a speech system; and

3 a speech input device connected to the speech system; wherein the
4 speech system comprises,

5 a voice server, wherein the server includes an integrated speech
6 and speaker recognizer that,

7 receives a first utterance from an intended talker via the speech
8 input device;

9 creates a voice characteristic model for the intended talker;

10 receives a second utterance from the intended talker via the
11 speech input device in a noisy environment;

processes a portion of speech associated with the second utterance, wherein the processor computes a speaker verification score with the voice characteristic model associated with the portion of speech, computes a speech recognition score associated with the portion of speech, and generates a combined score by combining the speaker verification score and the speech recognition score; and selects a best hypothesis associated with the second utterance and based upon the combined score.

8. The system of claim 7, wherein the speech input device comprises a cellular telephone, an analog telephone, a digital telephone, and a voice over internet protocol device.
9. The system of claim 7, wherein the portion of speech includes a word, a sentence, a syllable, or a frame.
10. The system of claim 7, wherein the server is further configured to alter a search path in a Viterbi search used by a speech recognizer.
11. An integrated speech and speaker recognition system comprising:
means for receiving a first utterance from the intended talker;

3 means for generating a voice characteristic model for the intended
 4 talker;
 5 means for receiving a second utterance from the intended talker in a
 6 noisy environment at the speaker recognition system;
 7 means for processing a portion of speech associated with the second
 8 utterance, wherein processing comprises,
 9 means for computing a speaker verification score with the voice
 10 characteristic model associated with the portion of speech,
 11 means for computing a speech recognition score associated with
 12 the portion of speech, and
 13 means for generating a combined score by combining the speaker
 14 verification score and the speech recognition score; and
 15 means for selecting a best hypothesis associated with the second
 16 utterance and based upon the combined score.

1 12. The system of claim 11, wherein the portion of speech includes a word, a
 2 sentence, a syllable, or a frame.

1 13. The system of claim 11, wherein the means for processing further
 2 comprises means for altering a search path in a Viterbi search used by a
 3 speech recognizer on the second utterance.

- 1 14. The system of claim 11, wherein the means for identifying an intended
2 talker comprises means for using hotword speech recognition to identify
3 the intended talker.
- 1 15. The system of claim 11, wherein the noisy environment includes
2 additional speakers, music, stationary and non-stationary noise.
- 1 16. The system of claim 11, wherein the voice characteristic model includes a
2 voice print, personal profile and linguistic characteristics.
- 1 17. A machine-readable medium having stored thereon a plurality of
2 instructions, said plurality of instructions when executed by a machine,
3 cause said machine to perform a process comprising:
4 receiving a first utterance from the intended talker at an integrated
5 speech and speaker recognition system;
6 generating a voice characteristic model for the intended talker;
7 receiving a second utterance from the intended talker in a noisy
8 environment at the speaker recognition system;
9 processing a portion of speech associated with the second utterance,
10 wherein processing comprises,
11 computing a speaker verification score with the voice characteristic
12 model associated with the portion of speech,
13 computing a speech recognition score associated with the portion
14 of speech, and

15 generating a combined score by combining the speaker verification
16 score and the speech recognition score; and
17 selecting a best hypothesis associated with the second utterance and
18 based upon the combined score.

1 18. The machine-readable medium of claim 17 wherein the portion of speech
2 includes a word, a sentence, a syllable, or a frame.

1 19. The machine-readable medium of claim 17, having stored thereon
2 additional instructions when processing a portion of speech, said
3 additional instructions when executed by a machine, cause said machine
4 to perform altering a search path in a Viterbi search used by a speech
5 recognizer.

1 20. The machine-readable medium of claim 17, having stored thereon
2 additional instructions when identifying a intended talker, said additional
3 instructions when executed by a machine, cause said machine to perform
4 using hotword speech recognition to identify the intended talker.

1

- 1 21. The machine-readable medium of claim 17, wherein the noisy
2 environment includes additional speakers, music, stationary and non-
3 stationary noise.
- 1 22. The machine-readable medium of claim 17, wherein the voice
2 characteristic model includes a voice print, personal profile and linguistic
3 characteristics.
- 1 23. A method comprising:
2 receiving an utterance from an intended talker at a speech recognition
3 system;
4 computing a speaker verification score with a voice characteristic
5 model associated and with the utterance;
6 computing a speech recognition score associated with the utterance;
7 and
8 selecting a best hypothesis associated with the utterance and based
9 on both the speaker verification score and the speech recognition
10 score.
- 1 24. The method of claim 23, wherein the voice characteristic model is
2 obtained from a voice model database.
- 1 25. The method of claim 23, wherein the voice characteristic model is
2 obtained from a first portion of the utterance.

1 26. A speech recognition system comprising:
 2 a speaker verifier;
 3 a speech recognizer connected to the speaker verifier; and
 4 an input device connected to the speaker verifier and speech recognizer,
 5 wherein the input device receives an utterance from an intended talker; and
 6 wherein the speech recognizer generates a recognition score associated with the
 7 utterance, the speaker verifier generates a speaker verification score associated
 8 with the utterance; and the recognition score is combined with the verification
 9 score to select a best hypothesis of the utterance.

1 27. The speech recognition system of claim 26, wherein the speech
 2 recognizer and speaker verifier are software entities residing on a speech
 3 server, and wherein the speech server comprises a processor, a bus
 4 connected to the processor, and memory connected to the bus that stores
 5 the software entities.

1 28. The speech recognition system of claim 27, further comprising a database
 2 connected to the speech server, wherein the database stores a voice
 3 characteristic model of the intended talker.